PROTEINS: Structure, Function, and Genetics 53:326–328 (2003)

STRUCTURE NOTE

Crystal Structure of tRNA (m1G37) Methyltransferase From *Aquifex aeolicus* at 2.6 Å Resolution: A Novel Methyltransferase Fold

Jinyu Liu,[2] Weiru Wang,[1] Dong Hae Shin,[2] Hisao Yokota,[1] Rosalind Kim,[1] and Sung-Hou Kim[1,2]*
[1]*Berkeley Structural Genomics Center, Lawrence Berkeley National Laboratory, Berkeley, California*
[2]*Department of Chemistry, University of California, Berkeley, California*

Introduction. The Berkeley Structural Genomics Center (BSGC) has focused on *Mycoplasma* as its structural genomics target organisms because of their compact genome size as well as their relevance to human and animal pathogenicity (http://www.strgen.org). The gene for transfer RNA (m1G37) methyltransferase of *Aquifex aeolicus* (GI number 2983865) is one of the structural genomics targets of BSGC that has been selected as a homolog of the *Mycoplasma pneumoniae* (MP) gene MPN183, trmD (http://www.strgen.org/status/mptargets.html). The protein structures coded by either gene are not known. It is one of the tRNA-modifying enzymes that catalyzes the transfer of methyl group from S-adenosyl-L-methionine (AdoMet) to guanosine at position 37, the nucleoside adjacent to and 3′ of the anticodon. This protein is important for the maintenance of the correct reading frame during translation.[1,2] In all organisms, the tRNA-reading codons CUN (Leu), CCN (Pro), and CGG (Arg) contain at position 37 1-methylguanosine (m1G37), and tRNA (m1G37) methyltransferases from members of all three phylogenetic domains show sequence similarities.[3] Currently, no structural information is available for this protein family. Here, we report the crystal structure of the enzyme tRNA (m1G37) methyltransferase from *Aquifex aeolicus* at 2.6 Å resolution. The structure reveals a novel methyltransferase fold distinctly different from those of the most common methyltransferases.

MATERIALS AND METHODS

Protein production. The trmD gene from *A. aeolicus* was amplified by PCR using *A. aeolicus* genomic DNA template and inserted into pHM6g expression vector, a derivative of pSKB3 (a gift from Steve Burley, Rockfeller University) having a 6-His-maltose binding protein (MBP)-tobacco etch virus (TEV) sites as an N-terminal fusion. Protein was expressed in *Escherichia coli* strain BL21(DE3)/pSJS1244 Star[4] upon induction with 0.5 mM isopropyl-β-D-thiogalactopyranoside. Selenomethionyl (Se-Met) protein was prepared according to the method of Doublie.[5]

For protein purification, 15.5 g of cell paste were resuspended in 135 mL 50 mM Hepes buffer, pH 7.5, and lysed by sonication. In a first purification step, NaCl concentration of the supernatant was adjusted to 300 mM before loading of the His-MBP-tagged target protein onto a cobalt affinity column (Talon resin; Clontech, Palo Alto, CA). The target protein was eluted in 50 mM Tris-HCl, pH 8.0, 200 mM NaCl, 300 mM imidazole. For MBP cleavage, the eluted protein solution was mixed with 360 μg mTEV and dialyzed into 50 mM Tris-HCl, pH 7.5, 1 mM EDTA, 1 mM DTT at 4°C (buffer A) overnight. The dialyzed protein was then heated to 70°C for 20 min followed by centrifugation at 10,000 g to remove precipitated nontarget proteins. For further purification to remove 6-His-MBP, the protein solution was then loaded onto a 5-mL HiTrapQ anion exchange column (Amersham Pharmacia, Piscataway, NJ) and eluted with a linear NaCl gradient in buffer A from 0 to 0.5 M NaCl in 20 column volumes. The target protein was in the flowthrough fraction and was concentrated to 7 mg/mL in 20 mM Tris-HCl, pH 7.5, 10 mM NaCl, 1 mM EDTA, 1 mM DTT for crystallization. The homogeneity of the protein was assessed by dynamic light scattering (DLS) experiment using DynaPro99 DLS machine (Protein Solutions, NJ).

Crystallization. Screening for the crystallization condition was done by the sparse matrix screening method.[6] A 1-μL protein solution was mixed with a 1-μL reservoir solution containing 0.1 M sodium citrate, pH = 5.5, 20% PEG 3000, and equilibrated over 1 mL of reservoir solution. Crystals grew within a few days at room temperature.

Data Collection. For data collection, crystals were transferred to harvesting buffer (reservoir solution + 5% PEG

Grant sponsor: National Institutes of Health; Grant number: GM 62412.

W. Wang's current address is Plexxikon, Inc., 91 Bolivar Drive, Berkeley, California.

*Correspondence to: Sung-Hou Kim, Berkeley Structural Genomics Center, Department of Chemistry, University of California, Berkeley, CA. E-mail: SHKim@cchem.berkeley.edu

Received 9 April 2003; Accepted 10 April 2003

TABLE I. Statistics of X-Ray Diffraction and Structure Refinement

Data set	Data collection	
	Peak	Remote
Wavelength (Å)	0.9793	0.9640
Resolution (Å)	20–2.6	20–2.6
Redundancy[a]	3.2 (3.0)	3.3 (3.0)
Unique reflections[a]	32111 (2883)	31974 (2874)
Completeness (%)[a]	97.6 (96.6)	98.2 (95.7)
I/σ[a]	15.1 (3.1)	13.7 (3.0)
R sym (%)[ab]	5.4 (37.9)	5.0 (44)
Crystal parameters and structure refinement		
Space group		C2
Cell dimensions		a = 153.4 Å, b = 96.1 Å, c = 57.4 Å, β = 96.2°
Vm (Å^3/dalton)		2.24
No. of water molecules		81
Resolution range of reflections used		20.0–2.6
Amplitude cutoff		0
R_{cryst} (%)[c]		27.9
R_{free} (%)[d]		30.0
RMSD from ideal bond length (Å)		0.018
RMSD from ideal bond angle (°)		1.97
Ramachandran plot		
Residues in most favored and additional allowed regions (%)[e]		97.3
Residues in disallowed regions (%)[e]		0

[a]The number in parenthesis is for the outer shell.
[b]$R_{sym} = {}_{hkl\ i} \mid I - \langle I \rangle \mid / {}_{hkl\ I}$.
[c]$R_{cryst} = {}_{h} \| F_{obs}(h) \mid - \mid F_c(h) \| / {}_{h} \mid F_{obs}(h) \mid$ for all data.
[d]R_{free} was calculated from 5.2% of structure factor amplitudes that were excluded from refinement.
[e]As defined in PROCHECK.

3000) for approximately 30 s and flash-frozen in liquid nitrogen. Diffraction data were collected at the Advanced Light Source (ALS) Beamline 5.0.2 at 100°K with a Quantum 4 charge-coupled device (CCD) detector and processed with HKL 2000.[7] Data statistics are summarized in Table I.

Structure Solution and Refinement. The positions of all nine selenium sites within the three monomers in an asymmetric unit (excluding the N-terminal methionines) were identified from MAD data using SOLVE.[8] Initial phases were greatly improved by density modification with multidomain NCS averaging (DMMULTI, CCP4 suite[9] program). For each monomer, 217 of the 257 amino acids were built into the resultant maps using the program O[10] by using the selenium sites as guides. No density was visible for the first four N-terminal amino acids, part of the loop region that connects N-terminal domain and C-terminal domain (residues 165–178), and the C-terminus 22 amino acids. Refinement of the initial model was performed with the program CNS[11] using the peak data with both phase and NCS restraints. Three runs of manual rebuilding were performed by using sigmaA-weighted 2Fo-Fc and Fo-Fc maps and Ramachandran analysis as guides. The final refinement was done with the TLS program in CCP4[9] suite followed by two runs of manual model rebuilding and addition of waters. A total of 81 water molecules were added to the final model. PROCHECK[9] analysis shows that the model has good overall geometry, with 97.3% residues in the most favored regions and additional allowed regions, and no residues falling in disallowed region of Ramachandran space (Table I).

Results and Discussion.

Structure. We determined the structure of the tRNA (m1G37) methyltransferase from *Aquifex aeolicus* to 2.6 Å resolution using the multiple-wavelength anomalous dispersion (MAD) technique (Table I). The monomer of the protein consists of two separate domains connected by a long partially disordered loop region (Fig. 1). The N-terminal domain consists of seven β-strands and seven α-helices (residues 4–164), arranged as a parallel β-sheet sandwiched by α-helices. The C-terminal domain (residues 179–235) consists of a long loop region and four short α-helices. In the crystal, each asymmetric unit contains three monomers. Two of the monomers in the asymmetric unit associate into a tight dimer through a twofold noncrystallographic axis (Fig. 1). Another monomer forms the same kind of dimer through the twofold crystallographic axis. The dimer has approximate dimensions of 72 × 45 × 45 Å. The structure suggests that the dimer may be the functional form for the following reasons: 1) the interactions between monomers are very extensive and tight as manifested by many nonpolar interactions commonly observed in biologically functional oligomer interfaces[12]; 2) the surface area buried by the dimer (1920 Å^2 for each monomer) is comparable to those from other oligomeric proteins and is larger than the area buried by typical crystal contacts[12]; 3) DLS data of the protein solution also suggest that the tRNA (m1G37) methyltransferase from A. aeolicus is oligomeric (dimer or trimer) state in solution (data not shown).

New Methyltransferase Fold. A structural similarity search, performed by the DALI[13] server, with the coordinates of N-terminal domain of the protein, indicated that the protein is structurally similar to protein mt0001, a hypothetical protein from *Methanobacterium Thermoautotrophicum* [*z* score 6.8, Protein Data Bank (PDB) accession code 1K3R] with root-mean-square deviation (RMSD) values of 3.0 Å for the superimposition of 110 Cα atoms. The sequence identity between the N-terminal domain of our protein and mt0001 is only 4%. The structural similarity search with the coordinates of the C-terminal domain did not identify any similar structures in the DALI data base with *z* score of ≥2, indicating that this domain is a novel structural fold. In all DALI search results with *z* score of ≥2, no classic Rossman-fold type methyltransferase structures (found in all hitherto known methyltransferases) are identified, suggesting the tRNA (m1G37) methyltransferase structure represents a novel methyltransferase fold. Although Dali search did not find any structural homologs other than the hypothetical protein mt0001, there is a recent protein structure of another novel methyltransferase, YibK, which has a unique "knotted" topology.[14] Our structure does not have the knotted topology, and RMSD between YibK and our structure is 3.2

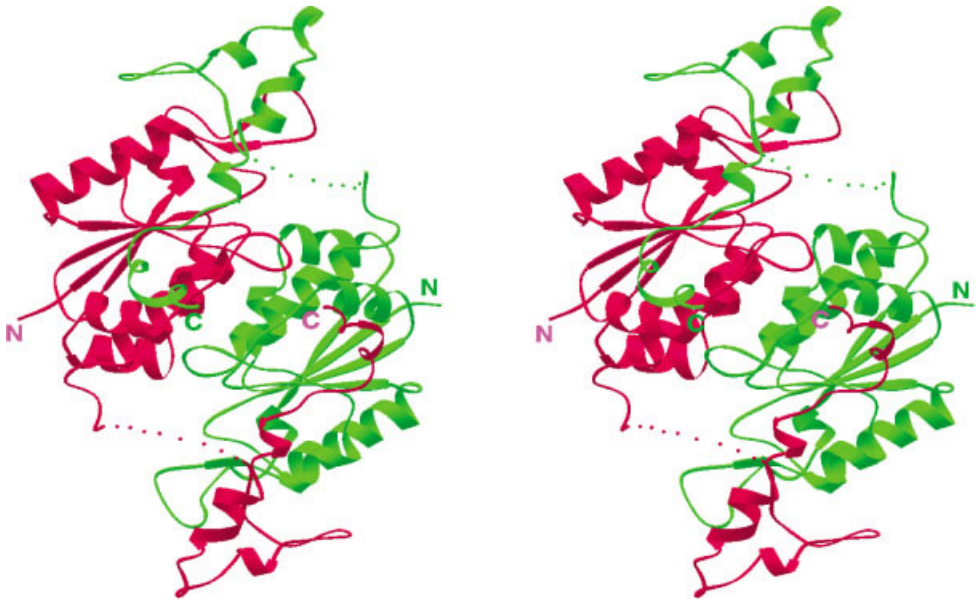

Fig. 1. Stereo ribbon diagram of the tRNA (m1G37) methyltransferase dimer structure. One subunit is drawn in green and the other in red. Produced with Ribbons[15].

Å for the best matched substructure of 91 Cα atoms out of about 164 Cα atoms. Our structure represents the first structure solved in the tRNA (m1G37) methyltransferase protein family and may be used as a new structural template to model the rest of the members of this protein family. The atomic coordinates of this structure was deposited in the Protein Data Bank under the accession codes 1OY5.

ACKNOWLEDGMENTS

We thank B. gold for cloning, J. Jancarik for help with the crystallization, and Diversa Corporation for kindly providing the genomic DNA of *Aquifex aeolicus*.

REFERENCES

1. Björk GR. Transfer RNA modification in different organisms. Chem Scr 1986;26B:91–95.
2. Hagervall TG, Tuohy TM, Atkins JF, Björk GR. Deficiency of 1-methylguanosine in tRNA from *Salmonella typhimurium* induces frame-shifting by quadruplet translocation. J Mol Biol 1993;232:756–765.
3. Björk GR, Jacobsson K, Nilsson K, Johansson MJO, Byström AS, Persson OP. A primordial tRNA modification required for the evolution of life? EMBO J 2001;20:231–239.
4. Kim R, Sandler SJ, Goldman S, Yokota H, Clark AJ, Kim S-H. Overexpression of archaeal proteins in *Escherichia coli*. Biotech Lett 1998;20:207–210.
5. Doublie S. Preparation of selenomethionyl proteins of phase determination. Methods Enzymol 1997;276:523–529.
6. Jancarik J, Kim S-H. Sparse matrixsampling: a screening method for crystallization of proteins. J Appl Crystallogr 191;24:401.
7. Otwinowski Z, Minor W. Methods Enzymol 1997;276:307–326
8. Terwilliger TC, Berendzen J. Automated structure solution for MIR and MAD. Acta Crystallogr 1999;D55:849–861.
9. Collaborative Computational Project No.4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr 1994;D50:760–763.
10. Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr 1991;A47:110–119.
11. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Gross-Kunstleve RW, Jiang J-S, Kuszewski J, Nilges N, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL. Crystallography and NMR system (CNS): a new software system for macromolecular structure determination. Acta Crystallogr 1998;D54:905–921.
12. Dasgupta S, Iyer GH, Bryant SH, Lawrence CE, Bell JA. Extent and nature of contacts between protein molecules in crystal lattices and between subunits of protein oligomers. Proteins 1997;28:494–514.
13. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol 1993;233:123–138.
14. Lim K, Zhang H, Tempczyk A, Krajewski W, Bonander N, Toedt J, Howard A, Eisenstein E, Herzberg O. Structure of the YibK methyltransferease from Haemophilus influenza (HI0766): a cofactor bound at a site formed by a knot. Proteins 2003;51:56–67.
15. Carson M. Ribbon models of macromolecules. J Mol Graphics 1987;5:103–106.